

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: _____

10/799,016 A

Source: _____

IFW16

Date Processed by STIC: _____

08/24/2006

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 08/24/2006

PATENT APPLICATION: US/10/799,016A

TIME: 12:57:43

Input Set : A:\-9-1-2.APP

Output Set: N:\CRF4\08232006\J799016A.raw

```

3 <110> APPLICANT: Gilbert, Michel
4      Wakarchuk, Warren W.
5      National Research Council of Canada
7 <120> TITLE OF INVENTION: Lipopolysaccharide alpha-2,3 Sialyltransferase of
8      Campylobacter jejuni and Its Uses
10 <130> FILE REFERENCE: 019633-000912US
12 <140> CURRENT APPLICATION NUMBER: US 10/799,016A
13 <141> CURRENT FILING DATE: 2004-03-11
15 <150> PRIOR APPLICATION NUMBER: US 60/078,891
16 <151> PRIOR FILING DATE: 1998-03-20
18 <150> PRIOR APPLICATION NUMBER: US 09/272,960
19 <151> PRIOR FILING DATE: 1999-03-18
21 <150> PRIOR APPLICATION NUMBER: US 10/058,636
22 <151> PRIOR FILING DATE: 2002-01-29
24 <160> NUMBER OF SEQ ID NOS: 8
26 <170> SOFTWARE: PatentIn Ver. 2.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 1293
30 <212> TYPE: DNA
31 <213> ORGANISM: Campylobacter jejuni
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (1)..(1293)
36 <223> OTHER INFORMATION: Campylobacter jejuni OH4384 cst-I gene
37      alpha2,3-sialyltransferase
39 <400> SEQUENCE: 1
40 atg aca agg act aga atg gaa aat gaa ctc att gtt agt aaa aat atg      48
41 Met Thr Arg Thr Arg Met Glu Asn Glu Leu Ile Val Ser Lys Asn Met
42 1      5      10      15
44 caa aat ata atc ata gca gga aat gga cct agc cta aaa aat att aat      96
45 Gln Asn Ile Ile Ile Ala Gly Asn Gly Pro Ser Leu Lys Asn Ile Asn
46      20      25      30
48 tat aaa aga ctg cct aga gaa tat gat gtt ttt agg tgt aac cag ttt      144
49 Tyr Lys Arg Leu Pro Arg Glu Tyr Asp Val Phe Arg Cys Asn Gln Phe
50      35      40      45
52 tat ttt gaa gat aag tat tat tta gga aaa aag att aaa gca gta ttt      192
53 Tyr Phe Glu Asp Lys Tyr Tyr Leu Gly Lys Lys Ile Lys Ala Val Phe
54      50      55      60
56 ttt aat cct ggt gtc ttt tta caa cag tat cac act gca aaa caa ctt      240
57 Phe Asn Pro Gly Val Phe Leu Gln Gln Tyr His Thr Ala Lys Gln Leu
58 65      70      75      80
60 ata cta aaa aat gag tat gaa ata aaa aat att ttt tgc tct aca ttt      288
61 Ile Leu Lys Asn Glu Tyr Glu Ile Lys Asn Ile Phe Cys Ser Thr Phe

```

RAW SEQUENCE LISTING

DATE: 08/24/2006

PATENT APPLICATION: US/10/799,016A

TIME: 12:57:43

Input Set : A:\-9-1-2.APP

Output Set: N:\CRF4\08232006\J799016A.raw

62	85	90	95	
64 aat tta cct ttt att gaa agc aat gat ttt tta cat caa ttt tat aat	336			
65 Asn Leu Pro Phe Ile Glu Ser Asn Asp Phe Leu His Gln Phe Tyr Asn				
66 100 105 110				
68 ttt ttc ccc gat gca aaa ctt ggc tat gaa gtt att gaa aac ctt aaa	384			
69 Phe Phe Pro Asp Ala Lys Leu Gly Tyr Glu Val Ile Glu Asn Leu Lys				
70 115 120 125				
72 gaa ttt tat gct tat ata aaa tac aat gaa att tat ttc aat aaa aga	432			
73 Glu Phe Tyr Ala Tyr Ile Lys Tyr Asn Glu Ile Tyr Phe Asn Lys Arg				
74 130 135 140				
76 att act tcg ggc gtc tat atg tgt gca att gct att gca tta gga tat	480			
77 Ile Thr Ser Gly Val Tyr Met Cys Ala Ile Ala Ile Ala Leu Gly Tyr				
78 145 150 155 160				
80 aaa acc atc tat tta tgt ggc att gat ttt tat gaa gga gat gtt att	528			
81 Lys Thr Ile Tyr Leu Cys Gly Ile Asp Phe Tyr Glu Gly Asp Val Ile				
82 165 170 175				
84 tat cct ttt gaa gct atg agt aca aat ata aaa aca atc ttt cct gga	576			
85 Tyr Pro Phe Glu Ala Met Ser Thr Asn Ile Lys Thr Ile Phe Pro Gly				
86 180 185 190				
88 ata aaa gat ttc aaa cct tca aat tgt cat tct aag gaa tac gat ata	624			
89 Ile Lys Asp Phe Lys Pro Ser Asn Cys His Ser Lys Glu Tyr Asp Ile				
90 195 200 205				
92 gaa gca tta aaa ttg tta aaa tca ata tac aaa gtt aat atc tac gca	672			
93 Glu Ala Leu Lys Leu Leu Lys Ser Ile Tyr Lys Val Asn Ile Tyr Ala				
94 210 215 220				
96 ttg tgt gat gat tct att ttg gca aat cat ttt cct tta tca att aat	720			
97 Leu Cys Asp Asp Ser Ile Leu Ala Asn His Phe Pro Leu Ser Ile Asn				
98 225 230 235 240				
100 att aat aac aat ttc act tta gaa aat aag cat aat aat tct ata aat	768			
101 Ile Asn Asn Asn Phe Thr Leu Glu Asn Lys His Asn Asn Ser Ile Asn				
102 245 250 255				
104 gat att tta ttg act gat aat act cct ggc gta agt ttt tat aaa aat	816			
105 Asp Ile Leu Leu Thr Asp Asn Thr Pro Gly Val Ser Phe Tyr Lys Asn				
106 260 265 270				
108 caa ctt aaa gct gat aat aaa att atg ctt aat ttt tat aat att ctt	864			
109 Gln Leu Lys Ala Asp Asn Lys Ile Met Leu Asn Phe Tyr Asn Ile Leu				
110 275 280 285				
112 cat tct aaa gat aat tta att aaa ttt tta aac aaa gaa att gcg gta	912			
113 His Ser Lys Asp Asn Leu Ile Lys Phe Leu Asn Lys Glu Ile Ala Val				
114 290 295 300				
116 tta aaa aaa caa acc act caa cga gct aaa gca aga atc caa aac cat	960			
117 Leu Lys Lys Gln Thr Thr Gln Arg Ala Lys Ala Arg Ile Gln Asn His				
118 305 310 315 320				
120 cta tcc tat aaa cta gga caa gct ttg att ata aat tct aaa agt gta	1008			
121 Leu Ser Tyr Lys Leu Gly Gln Ala Leu Ile Ile Asn Ser Lys Ser Val				
122 325 330 335				
124 tta ggt ttt tta tct tta cct ttt ata ata tta agt atc gtt att tca	1056			
125 Leu Gly Phe Leu Ser Leu Pro Phe Ile Ile Leu Ser Ile Val Ile Ser				
126 340 345 350				

RAW SEQUENCE LISTING

DATE: 08/24/2006

PATENT APPLICATION: US/10/799,016A

TIME: 12:57:43

Input Set : A:\-9-1-2.APP

Output Set: N:\CRF4\08232006\J799016A.raw

```

128 cat aaa caa gaa caa aag gct tat aaa ttt aaa gta aag aaa aat cca 1104
129 His Lys Gln Glu Gln Lys Ala Tyr Lys Phe Lys Val Lys Lys Asn Pro
130      355      360      365
132 aat tta gct tta cct cct tta gaa act tat cct gat tat aat gaa gct 1152
133 Asn Leu Ala Leu Pro Pro Leu Glu Thr Tyr Pro Asp Tyr Asn Glu Ala
134      370      375      380
136 tta aaa gaa aaa gaa tgt ttt act tat aaa tta gga gaa gaa ttt ata 1200
137 Leu Lys Glu Lys Glu Cys Phe Thr Tyr Lys Leu Gly Glu Glu Phe Ile
138 385      390      395      400
140 aaa gct ggt aag aat tgg tat ggg gag ggg tat atc aaa ttt ata ttc 1248
141 Lys Ala Gly Lys Asn Trp Tyr Gly Glu Gly Tyr Ile Lys Phe Ile Phe
142      405      410      415
144 aaa gat gtt cct agg ttg aag aga gag ttt gag aaa ggg gaa taa 1293
145 Lys Asp Val Pro Arg Leu Lys Arg Glu Phe Glu Lys Gly Glu
146      420      425      430
149 <210> SEQ ID NO: 2
150 <211> LENGTH: 430
151 <212> TYPE: PRT
152 <213> ORGANISM: Campylobacter jejuni
154 <400> SEQUENCE: 2
155 Met Thr Arg Thr Arg Met Glu Asn Glu Leu Ile Val Ser Lys Asn Met
156 1      5      10      15
158 Gln Asn Ile Ile Ile Ala Gly Asn Gly Pro Ser Leu Lys Asn Ile Asn
159      20      25      30
161 Tyr Lys Arg Leu Pro Arg Glu Tyr Asp Val Phe Arg Cys Asn Gln Phe
162      35      40      45
164 Tyr Phe Glu Asp Lys Tyr Tyr Leu Gly Lys Lys Ile Lys Ala Val Phe
165      50      55      60
167 Phe Asn Pro Gly Val Phe Leu Gln Gln Tyr His Thr Ala Lys Gln Leu
168 65      70      75      80
170 Ile Leu Lys Asn Glu Tyr Glu Ile Lys Asn Ile Phe Cys Ser Thr Phe
171      85      90      95
173 Asn Leu Pro Phe Ile Glu Ser Asn Asp Phe Leu His Gln Phe Tyr Asn
174      100      105      110
176 Phe Phe Pro Asp Ala Lys Leu Gly Tyr Glu Val Ile Glu Asn Leu Lys
177      115      120      125
179 Glu Phe Tyr Ala Tyr Ile Lys Tyr Asn Glu Ile Tyr Phe Asn Lys Arg
180      130      135      140
182 Ile Thr Ser Gly Val Tyr Met Cys Ala Ile Ala Ile Ala Leu Gly Tyr
183 145      150      155      160
185 Lys Thr Ile Tyr Leu Cys Gly Ile Asp Phe Tyr Glu Gly Asp Val Ile
186      165      170      175
188 Tyr Pro Phe Glu Ala Met Ser Thr Asn Ile Lys Thr Ile Phe Pro Gly
189      180      185      190
191 Ile Lys Asp Phe Lys Pro Ser Asn Cys His Ser Lys Glu Tyr Asp Ile
192      195      200      205
194 Glu Ala Leu Lys Leu Leu Lys Ser Ile Tyr Lys Val Asn Ile Tyr Ala
195      210      215      220
197 Leu Cys Asp Asp Ser Ile Leu Ala Asn His Phe Pro Leu Ser Ile Asn

```

RAW SEQUENCE LISTING

DATE: 08/24/2006

PATENT APPLICATION: US/10/799,016A

TIME: 12:57:43

Input Set : A:\-9-1-2.APP

Output Set: N:\CRF4\08232006\J799016A.raw

```

198 225          230          235          240
200 Ile Asn Asn Asn Phe Thr Leu Glu Asn Lys His Asn Asn Ser Ile Asn
201          245          250          255
203 Asp Ile Leu Leu Thr Asp Asn Thr Pro Gly Val Ser Phe Tyr Lys Asn
204          260          265          270
206 Gln Leu Lys Ala Asp Asn Lys Ile Met Leu Asn Phe Tyr Asn Ile Leu
207          275          280          285
209 His Ser Lys Asp Asn Leu Ile Lys Phe Leu Asn Lys Glu Ile Ala Val
210          290          295          300
212 Leu Lys Lys Gln Thr Thr Gln Arg Ala Lys Ala Arg Ile Gln Asn His
213 305          310          315          320
215 Leu Ser Tyr Lys Leu Gly Gln Ala Leu Ile Ile Asn Ser Lys Ser Val
216          325          330          335
218 Leu Gly Phe Leu Ser Leu Pro Phe Ile Ile Leu Ser Ile Val Ile Ser
219          340          345          350
221 His Lys Gln Glu Gln Lys Ala Tyr Lys Phe Lys Val Lys Lys Asn Pro
222          355          360          365
224 Asn Leu Ala Leu Pro Pro Leu Glu Thr Tyr Pro Asp Tyr Asn Glu Ala
225          370          375          380
227 Leu Lys Glu Lys Glu Cys Phe Thr Tyr Lys Leu Gly Glu Glu Phe Ile
228 385          390          395          400
230 Lys Ala Gly Lys Asn Trp Tyr Gly Glu Gly Tyr Ile Lys Phe Ile Phe
231          405          410          415
233 Lys Asp Val Pro Arg Leu Lys Arg Glu Phe Glu Lys Gly Glu
234          420          425          430
237 <210> SEQ ID NO: 3
238 <211> LENGTH: 41
239 <212> TYPE: DNA
240 <213> ORGANISM: Artificial Sequence
242 <220> FEATURE:
243 <223> OTHER INFORMATION: Description of Artificial Sequence:CJ18F 5' primer
245 <400> SEQUENCE: 3
246 cttaggaggt catatgacaa ggactagaat ggaaaatgaa c 41
249 <210> SEQ ID NO: 4
250 <211> LENGTH: 60
251 <212> TYPE: DNA
252 <213> ORGANISM: Artificial Sequence
254 <220> FEATURE:
255 <223> OTHER INFORMATION: Description of Artificial Sequence:CJ40R 3' primer
257 <400> SEQUENCE: 4
258 cctaggtcga ctcattagt gtgatggtgg tgatgttccc ctttctcaaa ctctctcttc 60
261 <210> SEQ ID NO: 5
262 <211> LENGTH: 231
263 <212> TYPE: PRT
264 <213> ORGANISM: Haemophilus influenzae
266 <220> FEATURE:
267 <223> OTHER INFORMATION: Haemophilus influenzae Rd putative open reading
268 frame (ORF)
270 <400> SEQUENCE: 5

```

RAW SEQUENCE LISTING

DATE: 08/24/2006

PATENT APPLICATION: US/10/799,016A

TIME: 12:57:43

Input Set : A:\-9-1-2.APP

Output Set: N:\CRF4\08232006\J799016A.raw

```

271 Met Gln Leu Ile Lys Asn Asn Glu Tyr Glu Tyr Ala Asp Ile Ile Leu
272   1           5           10           15
274 Ser Ser Phe Val Asn Leu Gly Asp Ser Glu Leu Lys Lys Ile Lys Asn
275           20           25           30
277 Val Gln Lys Leu Leu Thr Gln Val Asp Ile Gly His Tyr Tyr Leu Asn
278           35           40           45
280 Lys Leu Pro Ala Phe Asp Ala Tyr Leu Gln Tyr Asn Glu Leu Tyr Glu
281   50           55           60
283 Asn Lys Arg Ile Thr Ser Gly Val Tyr Met Cys Ala Val Ala Thr Val
284   65           70           75           80
286 Met Gly Tyr Lys Asp Leu Tyr Leu Thr Gly Ile Asp Phe Tyr Gln Glu
287           85           90           95
289 Lys Gly Asn Pro Tyr Ala Phe His His Gln Lys Glu Asn Ile Ile Lys
290           100          105          110
292 Leu Leu Pro Ser Phe Ser Gln Asn Lys Ser Gln Ser Asp Ile His Ser
293           115          120          125
295 Met Glu Tyr Asp Leu Asn Ala Leu Tyr Phe Leu Gln Lys His Tyr Gly
296           130          135          140
298 Val Asn Ile Tyr Cys Ile Ser Pro Glu Ser Pro Leu Cys Asn Tyr Phe
299 145           150          155          160
301 Pro Leu Ser Pro Leu Asn Asn Pro Ile Thr Phe Ile Leu Glu Glu Lys
302           165          170          175
304 Lys Asn Tyr Thr Gln Asp Ile Leu Ile Pro Pro Lys Phe Val Tyr Lys
305           180          185          190
307 Lys Ile Gly Ile Tyr Ser Lys Pro Arg Ile Tyr Gln Asn Leu Ile Phe
308           195          200          205
310 Arg Leu Ile Trp Asp Ile Leu Arg Leu Pro Asn Asp Ile Lys His Ala
311           210          215          220
313 Leu Lys Ser Arg Lys Trp Asp
314 225           230
317 <210> SEQ ID NO: 6
318 <211> LENGTH: 6
319 <212> TYPE: PRT
320 <213> ORGANISM: Artificial Sequence
322 <220> FEATURE:
323 <223> OTHER INFORMATION: Description of Artificial Sequence:6 His tail,
324      (His)6, polyhistidine tag, six adjacent histidines,
325      metal chelate affinity ligand
327 <400> SEQUENCE: 6
328 His His His His His His
329   1           5
332 <210> SEQ ID NO: 7
333 <211> LENGTH: 4
334 <212> TYPE: PRT
335 <213> ORGANISM: Artificial Sequence
337 <220> FEATURE:
338 <223> OTHER INFORMATION: Description of Artificial Sequence:linker between
339      two fusion protein domains
341 <400> SEQUENCE: 7

```

VERIFICATION SUMMARY

DATE: 08/24/2006

PATENT APPLICATION: US/10/799,016A

TIME: 12:57:44

Input Set : A:\-9-1-2.APP

Output Set: N:\CRF4\08232006\J799016A.raw